ECHO OPERATIONS PLAN

December 10, 2002

Version 1.0

Table of Contents

Preface

This document describes the policies and procedures required in the day-to-day decisions for operation and maintenance of the **EOS Clearing HOuse (ECHO)**.

This version 1.0 supersedes all previous versions.

This ECHO Operations Plan (EOP) has been developed and written by Patrick Agbu, Chao-Hsi Chang, Frank Corprew, Claudia Castaneda, Pooran Shukla, and Alex Lai, members of EOS Data Gateway Science and Operations task group.

Significant attendee input resulting from the ECHO Operations Workshop has been taken into consideration in formulating this document.

Acknowledgement

The authors would like to acknowledge the valuable comments provided by National Aeronautics and Space Administration (NASA) Earth Science Data and Information System (ESDIS) Project, Distributed Active Archive Center (DAAC) representatives, and other stake holders in ECHO that have led to numerous improvements in this plan.

We wish to thank Jennifer Farnham for meticulous proofreading and many useful suggestions.

1.0 Objectives

The Science and Operations Group will develop, implement and maintain an Operational Plan to support EOS Clearing House (ECHO). The plan will provide cohesive strategies and procedures needed to continue operations and execute a recovery in the event of an interruption that compromises the ability of the system to carry out its critical functions. On an annual basis, on or before June 30, the Science and Operations Group will submit to NASA/ESDIS a written certification that the ECHO Operations Plan has been reviewed, updated, and implemented according to protocol. The Group Manager will be responsible for all logistical planning, plan guidance, and plan maintenance.

2.0 Overview

ECHO functions as a metadata clearinghouse and order broker for ESDIS data and services. It hosts a cache of metadata representing the data holdings of a wide variety of providers. It adds value to existing systems by providing a single portal on the Internet where these metadata can be searched. The system provides an infrastructure that allows various communities to contribute metadata, services, and tools. As a metadata clearinghouse, it will support old and new data access paradigms such as navigation and discovery. As a data order broker it forwards orders for data discovered through the metadata search process to the data providers to fill. As a service broker, it decentralizes end user functionality and supports interoperability of distributed functions.

The system supports Data Providers, which are organizations that supply inventory-level metadata representing their data holdings. Operations staff will assist participating data providers to prepare and ingest their metadata, with minimum investment of resources and effort on the data provider's part.

The end user needs are addressed through a set of well-defined and open interfaces upon which the user community can build its own client applications. Specifically, the system supports extendable, flexible user interfaces, allowing industry and the community to contribute to the progress of available Earth Science applications. Groups outside the ESDIS community can also subscribe to metadata holdings in order to build their own systems. This approach allows users to build their own user interfaces (clients) to ECHO. For data providers, the system off-loads the burden of providing the resources required for searching and gives them the flexibility to support community-specific services and functionality. The system interoperability features allow all participants to benefit from the distributed development of functions, hence, reducing dependence on NASA resources.

3.0 ECHO Operations

The ECHO system is being developed to provide flexibility to NASA's EOS Data and Information System to better meet the needs of the science community. It is expected to be a mechanism to allow people to find out what data is available in a single place. The opportunities presented by having NASA's Earth Science metadata in one place go beyond just providing flexibility for accessing data, it is also expected that new functions and services not currently envisioned will be developed. A major feature of the system is that all interactions with it occur using the extensible markup language (XML) as the base message format. This gives the system an extensible base upon which to build.

The system supports both guest and registered users. Guests are non-persistent users of the ECHO system that can search for data, order products available to guests and download data products available online. But guest queries have no persistence in the system and guest access to order history is limited. Registered users have full search and order privileges, access to order status and history and their queries are maintained from session to session. Potential customers for the data and services published through the system can become registered users only through the auspices of an ECHO client. The Client supports all end user interfaces by formulating user queries, orders, etc. and passing them on to the ECHO system. When a Client receives a query or an order status results from the system, the Client formulates and displays the response to the user. In this context 'Client' means a software system utilizing ECHO; organization providing such systems to the science community is referred to as a 'Client Provider'.

3.1 Operational Roles

Different roles are assigned to groups in order to ensure the smooth operation of the ECHO system. The four major groups that play significant roles are: the ECHO Development Group (ECHO Dev), Data Providers, Client Providers, and the ECHO Operations and Maintenance Group (EOMG).

3.1.1 ECHO Development Group (ECHO Dev)

ECHO is a large software system developed by ECHO Dev, a group comprised of software design architects, software engineers and network communication and database specialists. This group will continue to provide valuable assistance to users in programming the Application Programming Interfaces (APIs) and to the EOMG in various other ways to provide assistance and advice in execution, maintenance, and updates of the ECHO system.

3.1.2 ECHO Operations and Maintenance Group (EOMG)

The EOMG as a group has been contracted by ESDIS to be responsible for the operation and upkeep of the hardware and software of the ECHO system. These responsibilities include:

- Database Administration – manage the Oracle database, maintain database tables and ensure account set up and provide catalog service support.
- Schedule Management – schedule and monitor the ingest of new metadata collections and metadata updates.
- Metadata Knowledgebase Supervision - validate the metadata to ensure that there are no errors and/or misunderstanding of metadata content.
- Valids Scientific Support – support the Data Provider in converting input metadata into an XML format and ensure that the metadata conform to standards and are mapped correctly to ECHO database formats.

3.1.3 Data Providers

Data Providers are those organizations that hold Earth science data and that provide the ECHO system with their metadata. These groups include but are not limited to entities such as Archives or Data Centers that manage science data, which can be subdivided into two groups as follows:

- Order Distribution Data Providers - these supply data to users in response to data order requests in various media including retrieval from URLs (Uniform Resource Locators).
- Non-order Distribution Data Providers - only supply metadata to the system and rely on metadata to describe how to retrieve data via URLs.

3.1.4 Client Providers

Client Providers are organizations that build client software to work with the ECHO system to search and retrieve metadata for their users.

- Client Developers - interact with ECHO Dev and EOMG to learn and apply the APIs, and with end-user communities to determine functional requirements of the Client.
- Client Maintenance and Client Operations Staff - collaborate with the EOMG to maintain communication between the Client and the ECHO system.

3.2 Specific Operational Functions

EOMG will serve as the hub for direct interactions between data providers and client providers and indirectly with client users. EOMG functions that will be performed are described in context of its interactions with these groups as indicated below.

3.2.1 Data Provider Interactions

These interactions involving various issues are summarized below.

3.2.1.1 Data Provider Acceptance and Registration

Initially, potential data providers will use the on-line website to establish contact with the EOMG. Commercial data providers are not supported in ECHO. The EOMG and the data provider will work together to establish an account via the ECHO Provider Registration Service. This service allows for registering Data Providers, establishing/updating Data Provider policy, establishing access controls to metadata, registering Client Providers, and delivering account status.

After an account has been successfully established, the prospective Data Provider gives a description of her/his metadata holdings and completes a Data Provider profile. The EOMG works with ESDIS to make a determination for approval or disapproval.

Flow of Events

a. The Data Provider contacts EOMG (e.g., email, URL) for Data Provider registration and provides information such as provider name, type of data, anticipated volume of metadata and submits an application to ECHO.

b. EOMG reviews the Data Provider information for completeness, accuracy, and acceptance status. With the exception of a non-acceptance status (denial of acceptance, or decline) the Data provider is granted approval and notified with a request for Data provider policy declarations.

   ALTERNATIVE1: EOMG rejects the application due to incomplete information contained within the registration document and sends a written notification of decision with accompanying rationale to the organization.

ALTERNATIVE2: EOMG rejects the application indicating that data is deemed inappropriate for the user community environment. Data Provider is notified of decision and reason.

c. The EOMG, upon accepting Data Provider, sets up all necessary system accounts and directories for use in Data Provider metadata pushes, and regularly scheduled ECHO processing.

d. The EOMG sets up necessary database schemas, tables, and functions for use in receiving Data Provider metadata.

e. The EOMG performs testing and monitoring of initial ingests from new Data Provider to insure proper processing and visibility in ECHO.

f. EOMG requires that the Data Provider conduct a series of tests before making their data holdings available to the user community. ECHO provides a test site and the EOMG supports the Data Provider in conducting all necessary tests to validate the Data Provider's metadata and the system-to-system interfaces.

g. Once testing has been successfully completed and access controls to the metadata have been defined, the Data Provider's data holdings are made available to the ECHO community.

Establishing a new data provider leads to increased load on the storage resources. Although the system is planned to be scalable, the resulting impacts are closely monitored. The EOMG will use this information to establish the ECHO system usage requirements for Data Providers. The next section outlines the basic features of such requirements.

3.2.1.2 Template Agreement for Data Provider

In order to ensure that the system provides the greatest level of service to the user community, and clients can appropriately describe the respective data holdings to the end users, it will be necessary to draw up an agreement with any Data Provider wishing to utilize the ECHO system. Some areas of agreement that may need to be considered, are outlined below:

a. The ability to provide and maintain a basic set of communications settings, so the user community, and/or the ECHO system can interact with the Data Provider.

b. To provide a basic set of acceptable metadata or acceptable metadata mapping, describing data holdings with respect to spatial

information, temporal information, search attributes, order options, etc.

c. To provide a schedule of metadata submission, so that system resources can be managed in the most efficient manner possible. This will eliminate potential adverse impacts to the ability of other Data Providers to use the ECHO system.

d. To provide support to the EOMG to resolve metadata problems. (such as, updates, metadata removal, etc.)

e. To provide support to the user community with regard to the data holdings made available for search and order via the ECHO system.

f. To provide support to the user community and/or EOMG with regard to resolution of order problems.

g. To manage accessibility to any data with restrictions based on the data, and/or the user community via access control lists.

h. To participate in the sanctioning of those Clients that will provide a search interface to Data Provider's data holdings.

Appendix C contains an example of the Data Provider Interface Control Forms.

3.2.1.3 Data Provider Policy Establishment/Update

The Data Provider Account Service allows registered Data Providers to set policy information and maintain basic organization and contact information. This function is automated and the Data Provider will have an API for establishing policy and making changes. The EOMG should only be required to step in when something is not functioning correctly, and to help Data Providers if they need assistance in defining required and optional functions.

Flow of Events

The Data Provider completes and submits to the EOMG a policy declaration. The Policy Declaration will have ECHO-defined defaults for handling of:

a. Price Quoting -- whether or not the Data Provider will quote a price on an order.

b. Maximum number of granules allowed for an order or maximum entries per order

c. Billing--payment or billing methods supported, (e.g. specific credit cards and purchase orders).

d. Distributed Search options (whether Data Provider supports)

e. Inventory Update

f. User Information.

ALTERNATIVE: Organization refuses to participate, based upon ECHO established framework and guidelines.

3.2.1.4 Access Control Management

The ECHO system provides an Access Control List (ACL) capability to control access to data as directed by the Data Providers. The Data Providers or their designated group member will submit an access control list containing identifiable groups of users for controlling access to collection as well as granule metadata. Access Control Management is an automated function that will be accomplished via two APIs: one to create and manage a named list of users, and another to establish which granules in a collection are hidden or restricted and which groups have access. The ACL takes effect immediately after the APIs are run. The EOMG has no ACL management responsibilities, but will be involved in troubleshooting problems.

Flow of Events

a. Data Provider organization completes and submits an ACL to ECHO. The ACL will contain:

- Named groups of users with accompanying rules and permissions for use in controlling access to metadata

  ALTERNATIVE1: Organization does not control access to metadata.

- Named groups of users with accompanying rules and permissions for use in controlling access to browse imagery (TBD)

ALTERNATIVE2: Organization does not control access to browse. (TBD)

- Named groups of users with accompanying rules and permissions for use in controlling access to Data Provider services

  ALTERNATIVE3: Organization does not control access to services.

b. The EOMG works with Data Provider to automate Data Provider ACL rules and permissions within the system.

3.2.1.5 Initial Metadata Ingest

Once the Data Provider has successfully tested his/her system against the ECHO system, and has received approval from the EOMG to become operational, the organization populates the service and Data Provider directories along with the inventory metadata.

Flow of Events

a. Data Provider declares service offerings and states whether they are dataset specific. Service data are loaded to ECHO

.

b. Data Provider tests site (ref: Data Provider Test Scenario)

c. Data Provider receives approval to turn site "on" (go operational).

   ALTERNATIVE1: Data Provider site has not received necessary approval for operational status and requires more testing.

d. ECHO flags new Data Provider holdings as being new

3.2.1.6 Metadata Update

Registered Data Providers submit routine updates to policy information, service directory, collection metadata and inventory metadata. These updates can occur in batch and/or interactive modes. The EOMG provides assistance as required.

The EOMG will set up an automatic script to periodically check for metadata update files. Once a file is detected, it is copied into another directory and processing begins. The system turns the file into a format suitable for direct ingestion into the database and then ingests it into a

temporary area. The data is checked for conditions that are known to cause problems, and then is moved into the actual data storage tables for that Data Provider. This new data is added to the database with the automatic adjustment of database indices excepting the spatial index.

The EOMG database manager on a periodic basis performs spatial index updates. Therefore, searches for the new data that do not include spatial constraints will find the new data almost immediately, but spatial searches will only find the data once the periodic update is run. Alternatively, the system can receive an email indicating a file is available for further action.

When the update is complete, ECHO will send an email to the Data Provider indicating how many data granules and data collections were successfully ingested and how many were rejected and why.

Flow of Events

a. Data Provider has a new collection that he/she wants to register with the system. Data Provider uses the Data Provider Update Tool to produce standardized metadata mapping/translation (for collection level)

   ALTERNATIVE: Interactive update (for policy and directory level)

b. Data Provider executes translation of collection level attribute names and values as well as granule level attributes and valids. Data Provider makes result available to the ECHO system import mechanism (policy driven: local or remote).

   ALTERNATIVE: No translation is needed.

c. Data Provider informs the system to perform assimilation of collection update. The system authenticates Data Provider's message.

   ALTERNATIVE: Authentication fails. Operation aborts. The system sends message to the Data Provider with failure notification.

d. ECHO gets data from Data Provider. This implies that the system knows the Data Provider's host location for placing data. The authentication of the message containing the location or the use of a previously authenticated location guarantees the source of the data.

ALTERNATIVE: The ECHO system polls local "landing area" where Data Provider placed data in a secure fashion. This implies that the Data Provider had sole rights to writing the data into a known location local to the system.

e. The ECHO system verifies translation information provided.

   ALTERNATIVE: Translation not valid. Import operation fails. The ECHO system notifies the Data Provider.

f. The ECHO system adds collections to catalog. Any processing anomalies are the responsibility of ECHO. The ECHO system provides a success notification to the Data Provider.

   ALTERNATIVE: Processing anomaly occurs (space exhausted, algorithmic failure). ECHO system recovers and notifies the affected participants.

g. The ECHO system performs interactive update of Data Provider/Data Provider policy information. This includes the changing of such items as point of contact, address, description, and phone number for the Data Provider. Policies can also be set for the Data Provider at either a collection or Data Provider level. The policies are hierarchical in nature; meaning, that if a policy is not set for a collection, then the Data Provider's policy is used. If the Data Provider provides no policy, then ECHO default policy is used.

   ALTERNATIVE: No update required.

h. The ECHO system imports new inventory granules (applying transformation if needed according to policy).

   ALTERNATIVE: The ECHO system updates some *existing* granules.

3.2.1.7 Metadata Mapping

ECHO Data Providers are required to supply metadata files in XML format for ingest into the system. If the Data Provider does not maintain the metadata in XML format, the Data Provider is expected to convert it into an XML format, preferably the format dictated by the ECHO DTD. If an XML version of the metadata exists but it does not match the ECHO DTD, then the ECHO metadata mapping tool can be used to define and perform the conversion required.

The EOMG assists the Data Providers with alternatives to the metadata mapping tool for creating XML files conforming to the ECHO DTD. Some alternatives are:

a. Data Provider loads attributes into a mapping definition tool. The tool displays Data Provider's attributes and valids and offers ECHO attributes/valids options and definitions for mapping. Data Provider selects ECHO attributes/valids that map. (Assumption: all mapping is performed one dataset at a time.)
b. Data Provider views metadata for own datasets and copies and edits the attributes/valids to be used.
c. Data Provider views metadata from other (external) datasets and copies attributes/valids to be used.

3.2.1.8 Metadata Reconciliation

In the event that a Data Provider submits conflicting valids, the EOMG, with guidance from the ECHO ETC group, will work with the Data Provider to find an acceptable substitute valid.

Periodic inventory reconciliation between ECHO holdings and Data Provider holdings will be necessary in order to implement any desired changes. ECHO will periodically produce a summary of granules and collections holdings in store for use by providers to make comparison and follow-up query and audit.

EOMG will maintain lists of acceptable valids, which will be constructed with NASA's Global Change Master Directory (GCMD) and EDG's lists of "Acceptable Valids" initially. In the event that new/additional keywords are needed, the Data Provider will work with EOMG to have them added to the ECHO lists; GCMD will be informed of the addition for their consideration to add these keywords to their directory. These acceptable lists of valids will be published by the EMOG on web-pages and will be accessible to Data and Client Providers.

3.2.1.9 System Resources Impact

The EOMG assesses the Data Provider needs for use of the system resources and further ensures that the Data Provider makes proper use of such resources within reasonable limits. If greater system resources are anticipated, these needs are brought to the EOMG so that justification for hardware upgrades can be made.

3.2.2 Client Provider Interactions

The Clients, being computer software that interface with ECHO via a client API, are the requesters of metadata from the ECHO system. The Client Providers interact with EOMG to learn and apply the ECHO Client API with particular end-user communities to determine functional requirements for their services. End users of Clients are customers of the Client Providers and not of ECHO.

3.2.2.1 Client Provider Acceptance

When an organization decides to build a Client to ECHO, they make a formal request to the EOMG providing a description of how they intend to work with the system, their customer profile information and the respective services they will offer their customers. The EOMG works with the prospective Client Providers to initialize their access to services and products and formalize a business relationship.

Flow of Events

a. Client Provider discovers ECHO and contacts the EOMG.
b. The Client Provider coordinates the development of Client to meet the needs of Data Providers and of the user community.
c. The EOMG assists Client Provider with requirements needed to utilize the API interface, as well as to gain the ECHO system and Data Provider sanctioning.
d. Client Provider builds Client, and notifies the EOMG of Client readiness.
e. The EOMG reviews and sanctions Client based on a set of developed requirements by the ECHO system and the Data Providers. Some requirements include
    - Client ability to identify and report problems
    - Client ability to resolve metadata problems
    - Client's adherence to reasonable use of ECHO resources.
f. The Data Provider will then sanction the Client based on set requirements.
g. If Client fails to meet the requirements for operation in the ECHO system, the EOMG decides if access should still be granted to user community at the users risk.
h. Client is operational and access is given to the user community.
i. Client operations are monitored for use of ECHO resources, ability to meet user needs, and adherence to established requirements. The inability of a Client to continue to meet user needs, adhere to requirements, or excessive use of the system resources at the

detriment of other Clients, will place Client under review for removal from the ECHO system.

3.2.2.2 Template Agreement for Client Provider

It will be necessary to draw up an agreement with any Client Provider wishing to use the ECHO system, to ensure that the ECHO system gives the greatest level of service to Data Providers and the user community. Anyone, including a commercial group, can become a Client Provider. Some areas of agreement that should be considered are:

a. Clients shall acknowledge use of ECHO by posting the ECHO partnership logo on their User Interface.

b. Use of suitable valid mapping between the Client and the ECHO system to insure a user utilizing a Client sends meaningful queries.

c. Clients will make use of constraints, available to the user, to better target queries. This action will keep query returns to a manageable size and ECHO system resource needs to a reasonable level without impacting the performance of other participating Client Providers.

b. Clients will establish Client User Services to assist end-users with data, order and data query issues.

e. Establish a process by which a Client can be removed from the ECHO system for reckless operation, and inability to provide adequate and appropriate support to its user community.

Appendix D contains an example of the Client Interface Control Form.

3.2.2.3 Client Maintenance and Operations Staff

Client Maintenance and Operations Staff will interact with EOMG to maintain communication between the Client and the ECHO system. The Client Operations Staff is responsible for interaction with end users.

3.2.2.4 ECHO End Users

The customers for the data and services published through the ECHO system, register through a Client. Clients support all end user interfaces with the ECHO system by formulating user queries, orders, etc. and passing them on to the system. When the Client receives query or order

status results from the ECHO system, the Client formulates/displays the response to the end user.

The ECHO end user registration process is automated and varies with the Client. ECHO provides an API for creating an account that the Client accesses to register users. The EOMG is not involved in the user registration process except to resolve problems that occur within this process and to provide education as needed.

3.2.2.5 Account Status Check

A user should be able to check the status of an account in terms of items ordered and items delivered through a client. The client should provide the option of requesting the cancellation of an order that has not yet been delivered.

The account status check process is automated in ECHO. The EOMG is not involved in this process except to respond to requests for account information not supported by the system or to clarify or correct information provided by the automated system. The central point of contact is echo@killians.gsfc.nasa.gov.

3.2.3 Service Providers

There is a role for service providers who add value to the ECHO data holdings by facilitating specialized services such as advanced search, subsetting, viewing, analyzing, measuring etc. for the data acquired by the users through ECHO. This may result in reducing the volume of data to be delivered enabling more complete and possibly faster science development. ECHO will encourage availability of a broad range of science tools and possible integration with data services. Service Providers are not operational now but will be given consideration in future in the same manner as the Client Providers.

3.2.4 Additional General Functions

3.2.4.1 Catalog Service Maintenance

The ECHO Catalog Service provides an interface used by Client Providers to search the system for science metadata. The Catalog Service provides the following automated functions:

a. Discovery - allows users to find out about the products offered through ECHO as well as information about the Data Providers.

b. Search for Metadata - enables the user, via a client-provided user interface, to formulate search criteria and issue a search that retrieves metadata that are then displayed to the user.

c. Browse – an extension to a search performed via an ECHO client in which the user can get a quick view of the data object to determine if it is of interest. In the event that the user knows the metadata, the browse data can be accessed directly.

ECHO responds automatically to user queries against the geo-spatial metadata database and provides a mechanism for Clients to convey the metadata that are available.

The aforementioned functions could also be available through a machine-to-machine interaction.

The EOMG has responsibility for maintaining the Catalog and the Catalog interfaces or APIs to the Clients.

3.2.4.2 Order Services Maintenance

Within ECHO, one is able to create an order and then add, delete, and update each item in the order as long as this is done before submitting the order to ECHO. Also within ECHO, one can look at the status and history of one's submitted and shipped orders.

The collection of catalog items that make up an order does not have to belong to just one provider, but can span many providers. In organizing providers and catalog items within an order, another concept called a 'provider order' is used. An order can consist of one or many provider orders. Each provider order can consist of one or many catalog items that belong to that same provider. When a full order is submitted to ECHO, it is these separate provider orders that are actually submitted to each associated provider.

The OrderEntryService API allows one to operate on orders, provider orders, and catalog items. All transactions within the OrderEntryService deal with orders before they are submitted to the system. Once the 'Submit' transaction is executed for a certain order within the OrderEntryService, the user can no longer execute any further changes on that order. However, a user can still monitor the current and historical status of any of user's submitted orders through the order-oriented transactions in the UserAccountService.

The transactions involved in OrderEntryService are:

a. AddOrderLineItem
b. CreateOrder
c. DeleteOrder
d. DeleteOrderLineItem
e. DeleteProviderOrder
f. ListUnsubmittedOrderSummary
g. PresentCatalogItem
h. PresentOptionDefinitionsForOrder
i. PresentOptionDefinitionsForProvider
j. PresentOrder
k. PresentProviderOrder
l. QuoteOrder
m. SetOptionSelectionsForOrder
n. SetOptionSelectionForProviderOrder
o. SetUserInformationForOrder
p. SubmitOrder
q. UpdateOrderLineItem
r. ValidateOrder

Once a provider order is submitted to the appropriate provider, the status of that order can be changed in two ways. First, the initial connection with the provider to send the provider order also allows for the provider to immediately send a response whether the submission of that order will be accepted or not. However, the provider can also use the ProviderOrderManagementService API which allows providers an asynchronous way of coming back to change the status of an order after they have had time to fully process the order.

The transactions involved in ProviderOrderManagementService are:

a. AcceptProviderOrderSubmission
b. CancelProviderOrder
c. ChangeTrackingID
d. CloseProviderOrder
e. PresentClosedOrder
f. PresentClosedOrderSummary
g. PresentOpenOrder
h. PresentOpenOrderSummary
i. PresentOpenOrderCancellation
j. RejectProviderOrderSubmission
k. RejectProviderQuote
l. SupplyProviderQuote
m. SupplyProviderQuote
n. UpdateStatusMessage

Wedged or unsatisfied orders will be detected by setting up query to look for orders that are not satisfied for more than a certain time period.

4.0 System Operations and Maintenance

4.1 System Hardware

The ECHO System hardware is contained in an environmentally controlled room.

Inventory of hardware components:

a. Two Sun 880 database servers with 4 CPUs and 733 MHz each, connected to 0.5TB hardware RAID
b. Two Sun E450 browse servers with 2 CPUs and 400MHz each.
c. For storage a 1.5 TB software RAID A5200 disk array.
d. Two Sun 880 web servers with 4 CPUs and 733 MHz each.
e. Two tape arrays, L20 and L4, each with two drives
f. An E220 backup server

4.2 System Monitoring

The EOMG monitors the general operating state of the system and performs designated routine tests to determine that specific application, network, server, and system software are functioning normally. The EOMG also responds to requests for investigation, diagnosis, and correction of system problems and performs all system level updates to system software and ensure hardware replacement.

4.3 Operational Constraints

The ECHO System is planned to be a 24-hour by 7 days per week operational system with redundant platforms for fail-over capability. This implies that the system is never turned off except when necessary to manage system outages. ECHO operations and system administration staff will be available for 8 hours per day (generally 9 am to 5 pm) 5 days per week (Monday through Friday) excluding holidays. Any system outages will be handled as a priority and will be dealt with in an expedient manner. In general, a hardware/software failure will be diagnosed and serviced within 8 hours. Hence such an event can temporarily disable the system for a day if the failure occurred early in the day. Otherwise the outage can continue over to the next working day. Such a situation is expected to occur rarely and the overall system availability is anticipated to approach 0.99. The system recovery plan is described in greater detail in Section 4.5.

4.4 Database Configuration and Data Components

The database is made up of a two-node Real Applications Oracle cluster on the Sun 880 servers with the two Sun E450 servers handling the browse data. The components are described below.

a. ECHO Database: the database is made up of the business schema and Data Provider schemas, each with varying data characteristics. For instance, there may be largely dormant providers while others may have large amounts of data coming in frequently.

b. ECHO Browse: This image file will experience relatively infrequent change but frequent additions.

c. ECHO Data Provider data ingest files: Data files for the data provider collections and granules.

d. ECHO Log files: The ingest logs generated while data is being ingested into the system.

e. Software Components: The ORACLE 9i database is the main software component of the system.

4.5 System Backup, Recovery and Security

The EOMG execute routine system backup functions to copy the information from the system machines, either the entire or partial system, for safe keeping for a specific time period. Routine system restore functions are executed to return the data to the machines to allow operations to continue from a specific point in time.

The ECHO system will use Oracle's 'Recovery Manager' feature to perform automatic physical and logical backup of the data and its recovery.

Alternative backup procedures shall include the use of scripts developed by the EOMG to physically copy the data files and archive the redo logs.

As the result of an event that destroys the system or integrity of the database, a complete system restore process is performed.

The process required to complete a successful restoration will be clearly documented. Tests to recover from known failures and disasters shall be executed on regular basis.

4.5.1 System Backup

Data will be prioritized depending on its importance and the degree to which it changes. The list below outlines the process for the System, Database, Browse Data and Ingest Files.

System

The system has a 3 level backup that covers the ECHO Software and the Configuration files for Unix and ECHO:

Full back up on the 1st of each the month with a retention period of 6 months. Archival will be conducted off site.
Each Friday, a weekly incremental back up will be completed and will include incremental backup that dates back to the last full back up.
Daily incremental back up will be conducted. Backup dates to the previous weekly incremental.

Database

Hot backups of the provider schemas, the business schema, and Control files.

Cold backups of the provider schemas, the business schema , the archived re-do logs, and Control files.

Browse Data

Full backup once a month and daily incremental

Ingest files

Permanent backup of ingest files to tape will be conducted and all provider metadata files and log ingest files will be deleted from disk. Currently, this process will be completed with each monthly full system backup.

4.5.2 System Recovery

4.5.2.1 Recovery Plan

In the event of a natural or human caused disaster, a recovery plan will be implemented to support the operations team in response to the event of a natural or human-caused disaster. The recovery plan must be reviewed and tested periodically (at least every two years or upon significant change). A copy of the plan should be kept at a location away from the system (at a minimum in another building), usually with the backup

materials, in case it is not possible to return to the facility. The following table includes all the hardware under the control of the operations team.

Hardware	Primary Support Staff	Backup Support Staff	Estimated restore time
Database Server	Staff A	Staff B	48 hours
Web Server	Staff C	Staff D	24 hours
File Server	Staff E	Staff F	12 hours
Backup Server	Staff G	Staff H	12 hours

4.5.2.2 Recovery Time

There are two types of recovery processes that may need to be carried out at any one time:

Instance recovery - relates to hardware failure. How long would it take for the database to open and be accessible again? There is fast start fault recovery functionality in Oracle9i, which reduces the time needed for cache recovery and makes it more predictable.

Media recovery - relates to events when a database file needs to be restored. Depends on the backup media, the time to restore the file and the application of transactions for that file. It is very important to run the database in ARCHIVELOG mode to enable full instance recovery.

4.5.2.3 Contingency Plan

Those responsible for managing the applications that run on the system (and the data owner) must plan ways that their data or application will continue performing critical functions if the facility in which processing normally occurs suddenly stops supporting the application. Due to different recovery requirements, a separate Contingency Plan may be required for each application.

Application	Primary Support Staff	Backup Support Staff	Estimated restore time
Application1	Staff A	Staff B	12 hours
Application2	Staff C	Staff D	12 hours
Application3	Staff E	Staff F	12 hours
Application4	Staff G	Staff H	12 hours

4.5.3 System Security

All the system security considerations should follow the NASA security procedures guideline document NPG 2810.1. Some of the important items to be implemented are listed below:

- The ECHO system allows a user to search and read documentation to determine if the data set under consideration is the one being sought. The data processed by and stored within the ECHO system is not considered to have confidentiality issues. The integrity of data processed and stored within the ECHO system is of prime concern.

- A privileged user is one who can alter or circumvent operating system security protections. This applies to users who may have only limited privileges, such as developers, but who can still bypass security precautions. Assigned privileges could result in a user having capability to modify system configurations, account privileges, audit logs, data files or applications. It is important to periodically check and verify if there is a continued need for the access level and privileges that are assigned. A list of privileged and limited privileged users will be maintained and a semiannual verification of the list will be carried out as part of the security procedures.

- A process will be implemented to ensure that userIDs are validated annually.

- Configuration will permit no more than five successive unsuccessful logon attempts.

4.6 Performance Monitoring

The following procedures will be used in performance monitoring:

- Routine polling of system-by-system software

- Monitoring of system health and status

- If the system indicates an anomaly or problem, the EOMG shall perform an analysis to determine the nature of the problem. If the analysis indicates that the problem is software related, the EOMG will take the necessary action to correct the problem.

- If the system indicates an anomaly or problem, the EOMG shall perform an analysis to determine the nature of the problem. If the analysis indicates that the problem is hardware related, the EOMG will take the necessary action to replace the defective hardware.

- Execution of routine application level tests will be performed to determine:
  If the provider software and links are active or
  If provider is able to order and access the system or
  Network bandwidth

- Monitoring of tests that are automatically scheduled and executed by the software.

- Upon receipt of a notice (email or phone call) from an ECHO user or Data Provider reporting a problem, the operations team acknowledges the problem report, investigates, and takes corrective action

- Development of statistics on system usage and system response time: Statistics will be collected on searches, domain traffic, client traffic, user characterization, accounts etc. Specifically, this will include

  i) Query Statistics: Minimum, Maximum, Average, Median time to complete a query and Minimum, Maximum, Average, Median number of granules returned in query
  ii) Order Statistics: Number of granules per order, Megabytes per order, Media type/electronic
  iii) Data Provider Statistics: Number of collections, granules ingested per day, week, month in ECHO, archive totals
  iv) Client Statistics: Number of queries, orders

EOMG and ECHO dev team will investigate commercial/shareware tools for Oracle, Weblogic and Solaris monitoring. Oracle Enterprise Manager may provide some of the required data. (TBD)

5.0 Training

The Science and Operations Manager will assure that training in all facets of ECHO operations is provided on an on-going basis to all appropriate personnel. Requisite training will enable all appropriate staff members to be adequately trained to fulfill their responsibilities in support of the functions and system recovery process. Training must include familiarity

with and a working knowledge of the ECHO Operations Plan. Training for new employees will be carried out within 30 days of start of work.

6.0 Change Management

Change Management exists to coordinate and inform customers of all changes that impact any shared computing systems or services under the direction of the ECHO Science & Operations manager.

6.1 The objectives of change management

a. To allow changes while, at the same time, maintain or improve service stability and availability in order to increase the probability of success.

b. To ensure that all parties affected are informed as early as possible of the planned changes.

c. To provide a record of changes implemented to assist with and shorten problem determination time.

d. To ensure that technical and management accountabilities for all changes are identified.

e. To assist with the accuracy of predictions of impact, such as response time, utilization.

f. To ensure that all affected parties are not only informed, but are also provided necessary documentation, and training is in place prior to the implementation of the change.

6.2 Change Process

To achieve the objectives of all levels of managers within the development and operation team the following order of procedures must be enforced:

a. Request for change

b. Change approval

c. Identify impact from the change

d. Schedule change

e. Notification of change schedule

f. Backup system

g. Implement change

h. Test change

i. Notification of change

6.3 Change Management Tools

A version control tool keeps track of all the changes for all the source code within the ECHO system. The name of the version control tool is TBD. Throughout the change process, the status of all changes will be tracked in the Access Database. The database must include the following columns:

a. System to be changed

b. Change Request Submitted by and date

c. Change Implemented by and date

d. Change Impact (Low, Medium, High)

e. Change Impact to Parties

f. Require Shut Down Database (yes/no)

g. Description of Change (summary and detail steps)

h. Schedule Date/Time for change

i. Approval Signature (Government and Contractor)

j. Implementation Status

k. Testing Result

l. Notification Status

7.0 Summary

This document summarizes important issues related to ECHO Operations. The EOMG is likely to refer to this document first when background information is required. In order to provide them with the essential material in one place we have included the background material with which future workers may not be familiar. Hence the document delineates

the metadata model in some detail in Appendix A. Appendix B provides the glossary/acronyms used in this document. Appendix C and D outline the basic information that will be required from a new Data Provider and Client Provider respectively. In compiling this document we have relied mostly on the material available at the ECHO web site (http://dangermouse.gst.com/ECHO/), ECHO Requirements and the presentations at the ECHO Operations Workshop.

Appendix A Metadata Data Model

The ECHO Data Model organizes and describes the metadata for the system. The data model consists of a bounded set of attributes intended to cover the essential characteristics of all earth science data sets. More details can be referenced at the ECHO website, http://dangermouse.gst.com/ECHO/ .

A.1 Metadata Terms

The following are some of the ECHO Metadata terms and their definitions:

Attribute: An attribute is the basic entity of a data model and is label for a type of information. An attribute is said to be populated when a data value is assigned to it.

Metadata: This term is used to define all descriptive information that accompanies data products from the providers. Metadata provides essential information about data products, which in turn can facilitate search, processing, distribution and other services required by users.

Core Metadata: ECHO Core metadata is a subset of attributes intended to cover the essential characteristics of all data sets in the ECHO system.

Mandatory Metadata: Certain core metadata attributes are considered mandatory, i.e. they must be supplied in order for the associated granule or collection to be accepted into the ECHO system.

Optional Metadata: These attributes are not required for every data product in the ECHO system and are optional. If supplied, these will enhance the search services available to users for the data product.

Product Specific Metadata: These are additional attributes associated with data products that are not contained in the ECHO metadata attributes. Product Specific Attributes (PSAs) may be held external to the data model and can be used for searches of data products. PSAs usually describe the characteristics of the data products such as the information describing specific characteristics of the instrument at the time of sensing or information that applies to a certain discipline.

A.2 Organization of the Metadata Model Objects

The ECHO Metadata Model was developed based on the EOSDIS Core System (ECS) data model. The metadata are clustered around two main entities and can be divided into two broad logical categories. They are:

Collection: Contains tables to define collection level metadata attributes. A collection is any logical grouping of logical or physical granules chosen by the data provider for identification, grouping, and advertisement in ECHO system as a collection.

Granule: Primary module of the metadata describing the data granule which is the smallest aggregation of data that is independently managed.

Other metadata describing spatial, temporal, contact, packaging and document information are populated at both collection and granule levels to properly describe data.

Total number of entities: 92
Total number of attribute: 458
Each entity's specification include:
Description
Annotation
Identifier list
Relationship list
Attribute list

Each attribute's specification (commonly called a "data dictionary") include:

Description
Content Source
Alias
Domain Value (valids)
Domain Description
Data Type
Units
Default Value
Format
Constraints (dependencies, conditions)

A.3 Core Attributes

DATASET_ID, CAMPAIGN_SHORT_NAME,
INSTRUMENT_SHORT_NAME,
PLATFORM_SHORT_NAME, SENSOR_SHORT_NAME,
DISCIPLINE_KEYWORD_ID, GRANULE_ID,
PROCESSING_LEVEL_ID are the core attributes of the ECHO system.

A.4 Modification of the Metadata Data Model

The ECHO Technical Committee (ETC) will make the final decision to accept/deny any data model modification request submitted to the committee.

Appendix B Glossary/Acronyms

Acronym	Expansion
ACL	Access Control List
API	Application Programming Interface
DAAC	Distributed Active Archive Center
DTD	Document Type Definition
EOS	Earth Observing System
EDG	EOS Data Gateway
EOSDIS	EOS Data and Information System
ECS	EOSDIS Core System
ECHO	EOS Clearing HOuse
ECHO Dev	ECHO Development Group
EOP	ECHO Operations Plan
EOMG	ECHO Operations and Maintenance Group
ETC	ECHO Technical Committee
ESDIS	Earth Science Data and Information System
GCMD	Global Change Master Directory
NASA	National Aeronautics and Space Administration
PSA	Product Specific Attribute
URL	Uniform Resource Locator
XML	Extensible Markup Language

Appendix C Provider and Collection Interface Control Forms (Generic)

Provider Interface Control Form

Provider Name:

_____

Provider Short Name: _____

Provider Type (Metadata, <Service>, <Both>, <Search>)

Number of Collections: _____

Type of metadata: (Image/large file metadata, metadata is data, Small file metadata, varies)

Desired frequency of update: (Hourly, Daily, Weekly, Monthly, Other, varies)

Desired update schedule (e.g. daily at 2AM, hourly at 7 minutes past the hour, weekly on Sundays at 6AM):

_____

Metadata mapping: (at provider, at ECHO, 3rd party, varies)

Contact Information for Provider:

Name: _____

Email: _____

Telephone: _____ Fax: _____

Address: _____

Addess2: _____

City: _____ State: _____ Zip: _____

Collection Interface Control Form (Generic)

Collection Name:

_____

Collection Short Name:_____

Provider: _____

Collection Type (Metadata, <Service>, <Both>, <Search>)

Number of granules:

currently: _____

expected to add: _____per_____occurring

how often:_____

expected to change:_____per _____occurring

how often: _____

expected to delete: _____per _____occurring

how often: _____

Approximate size of metadata in kilobytes per granule:

_____

Browse image available? (Yes / No) Average browse size:

_____

Type of metadata: (Image/large file metadata, metadata is data, Small file metadata)

Desired frequency of update: (Hourly, Daily, Weekly, Monthly, Other, varies)

Desired update schedule (e.g. daily at 2AM, hourly at 7 minutes past the hour, weekly on Sundays at 6AM):

_____

Metadata mapping: (at provider, at ECHO, 3rd party, varies)

Contact Information for Collection:

Name: _____

Email: _____

Telephone: _____ Fax: _____

Address: _____

Addess2: _____

City: _____ State: _____ Zip: _____

Appendix D Client Interface Control Form (Generic)

Client Name:

_____

Client Short Name: _____

Client Type (Metadata, <Service>, <Both>, <Search>, <Other>)

Number of Users: _____

Type of metadata: (Image/large file metadata, metadata is data, Small file metadata, varies (multiple/other)

Contact Information for Client:

Name: _____

Email: _____

Telephone: _____ Fax: _____

Address: _____

Addess2: _____

City: _____ State: _____ Zip: _____